

Venues in Toronto and NY

The Problem

Determine which city is the best to open new stores and which one is the best to invest in already existing ones.

Data

The data we will use is the following:

- name of the neighborhoods;
- geographical coordinates (latitude and longitude) of the neighborhood;
- 100 venues around the geographical coordinates of neighborhoods.

The data of the first item is obtained from Wikipedia or other sources. The data of the second item is obtained by the use of geocoder data or other sources. The data of the third item is obtained by the use of Foursquare location data.

Methodology

We process and prepare data of the financial districts of the cities and their surroundings. Due to the amount of people visiting the financial district, we can study all these location at the same time to understand the preferences and opportunities for each city.

We get the top 100 venues per neighborhood in a 500 meters radius, a bounded distance for people looking for a coffee or bakery.

We cluster the neighborhoods base on the similarities of their venues From this clusters we can infer how homogeneous are the cities.

This provide a better understanding of the neighborhoods and their interactions as elements of the same cluster.

Toronto Clusters

As a result of clustering the neighborhoods, we can obtain a big picture of the venues in Toronto by study only the red cluster.

The other clusters are very small containing at most one neighborhood. Therefore the neighborhoods that are not in the red cluster requires a study as an isolation neighborhood.

Toronto Isolated Clusters

Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Moore Park, Summerhill East	1	Playground	Tennis Court	Women's Store	Dance Studio	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run

```
----Moore Park, Summerhill East----
      venue  freq
0     Playground  0.5
1     Tennis Court  0.5
2     Yoga Studio  0.0
3     Movie Theater  0.0
4           Market  0.0
```

As we can see in this cluster we only have playgrounds and Tennis Courts. Therefore, this neighborhood is about physical activities and open a store about sport equipment is a good idea.

Toronto Isolated Clusters

Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Rosedale	2	Park	Playground	Trail	Dance Studio	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run

```
----Rosedale----
      venue  freq
0      Park  0.50
1  Playground 0.25
2      Trail  0.25
3  Yoga Studio 0.00
4  Movie Theater 0.00
```

As we can see in this cluster we only have playgrounds and parks. Therefore it is a family neighborhood with a kids friendly point of view, open a toy store would be a good idea.

Toronto Isolated Clusters

Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Forest Hill North & West	3	Jewelry Store	Trail	Mexican Restaurant	Sushi Restaurant	Women's Store	Department Store	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop

```
----Forest Hill North & West----
      venue  freq
0      Jewelry Store 0.25
1          Trail 0.25
2  Mexican Restaurant 0.25
3      Sushi Restaurant 0.25
4          Yoga Studio 0.00
```

This cluster jewelry plus restaurants cluster, therefore a wine store fits the needs of the cluster.

Toronto Isolated Clusters

Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Lawrence Park	4	Park	Swim School	Bus Line	Falafel Restaurant	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run

----Lawrence Park----

	venue	freq
0	Park	0.33
1	Bus Line	0.33
2	Swim School	0.33
3	Yoga Studio	0.00
4	Moroccan Restaurant	0.00

As we can see in this cluster we only have parks and swimming pools. Therefore, this neighborhood is about physical activities and open a store about sport equipment is a good idea.

Manhattan clusters

As a result we see that there are no isolated neighborhoods in the sense of similarities, but there are neighborhoods which are isolated in the sense of distance to the other elements of the cluster.

Due to this particularities, Manhattan is not homogeneous and therefore we can not study Manhattan by only study one of the clusters.

Even more, decision on the red cluster need a more deep study depending on the neighborhood (there is a red neighborhood very far from the other red neighborhoods).

Conclusions

Toronto is a more homogeneous city in terms of venues, and with 232 different kind of venues it is a more attractive city to investors in new kind of venues

New York, which is more heterogeneous and with 321 different kind of venues the opportunities for new venues are smaller.

Generally speaking new kind of venues have more possibilities of success in Toronto than in Manhattan. On the other hand due to the homogeneity of the city, the location of a venue is not an important factor at the moment of deciding which new service a bakery or cafe could offer in Toronto.

Clearly Manhattan is on the other side of the spectrum, it would be harder to start a new venue in Manhattan, but the data tells us that people commute more in Manhattan due to the heterogeneity of the area and therefore it is more convenient to invest on small stores to offer new services than to invest in new venues

To decide which new service to offer and where in Manhattan or what kind of new venue to open in Toronto, a more detail analysis is required, depending on the ambitions and goals of the investment.